Research Article

A Hybrid SVM and PSO Approach in The Classification of Hypertension at Medika Palopo General Hospital

[1,*] Rizka Utami, [2]Desi Anggreani, [3]Darniati

[1,2,3] Muhammadiyah University of Makassar, Rappocini sub-district, Makassar city, South Sulawesi 90221, Indonesia

* Corresponding Author: desianggreani@unismuh.ac.id

Abstract: Hypertension is a chronic disease that often goes undetected in its early stages, increasing the risk of complications such as stroke and heart failure. Accurate classification of hypertensive patients is essential to support early intervention and reduce morbidity and mortality. This study aims to evaluate the performance of the Support Vector Machine (SVM) algorithm and to develop a hybrid classification model by integrating Particle Swarm Optimization (PSO) to improve the predictive performance of SVM. The research was conducted using 400 patient records from Medika Palopo General Hospital, equally divided into hypertensive and non-hypertensive groups, with 12 clinical features as input variables. The SVM model was built using a sigmoid kernel with default parameters (C = 1.0, gamma = auto), while the hybrid model utilized PSO to optimize the values of C and gamma. Evaluation results show that the conventional SVM model achieved an accuracy of 61.25%, precision of 63.22%, recall of 56.50%, F1-score of 59.31%, and AUC of 0.6400. After optimization using PSO, the hybrid model significantly improved with an accuracy of 96.75%, precision of 97.14%, recall of 96.31%, F1-score of 96.73%, and AUC of 0.9725. The findings indicate that the hybrid SVM-PSO approach effectively enhances the classification performance of hypertension prediction models and offers promising potential to be developed into an AI-based medical decision support system.

Keywords: hypertension, artificial intelligence, classification, support vector machine, particle swarm optimization, machine learning

Citation: Utami, R., Anggreani, D., & Darniati. (2025). Review of navigation systems used for visually impaired individuals: Hybrid SVM and PSO approach in the classification of hypertension at Medika Palopo General Hospital. *Iota*, 5(3). https://doi.org/10.31763/iota.v5i3.1003

Academic Editor: Adi, P.D.P

Received: June 14, 2025

Accepted: July 15, 2025

Published: August 01, 2025

Publisher's Note: ASCEE stays neutral about jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Hypertension is a chronic condition that significantly contributes to cardiovascular morbidity and mortality. Conventional diagnosis methods, which rely on manual clinical assessments, often lack efficiency and scalability. This condition can lead to serious complications such as stroke, heart failure, myocardial infarction, kidney failure, and vision impairment [1]. According to the World Health Organization (WHO), more than 1 billion people worldwide suffer from hypertension, and the number continues to rise due to lifestyle changes and other risk factors. In Indonesia, the 2023 Indonesian Health Survey (SKI) [2] reported a prevalence of 30.8% among adults aged 18 and older. In Palopo, hypertension is among the most frequently treated conditions, with hospital data showing that most inpatients, especially pre-elderly and elderly patients, are diagnosed with hypertension [3].

The conventional classification of hypertension depends on clinical observations and laboratory tests, which are time-consuming and resource-intensive. In the era of artificial intelligence (AI) and data science, machine learning offers promising solutions to improve diagnostic efficiency and accuracy [4]. One of the most widely used and effective methods is the Support Vector Machine (SVM), which classifies data by identifying the optimal separating hyperplane. Research by [5] comparing SVM and neural networks in heart disease classification found SVM achieved 83% accuracy, outperforming neural networks (82%). However, SVM's performance depends heavily on selecting optimal parameters such as cost (C) and gamma (γ) in the RBF kernel [6]. Poor parameter selection can reduce accuracy significantly.

Particle Swarm Optimization (PSO) is an effective metaheuristic for optimizing SVM parameters [7]. Compared with Genetic Algorithms (GA), PSO offers faster convergence, fewer iterations, and simpler computation, as it does not require crossover or mutation operations [8]. Several studies have shown that PSO improves SVM performance in disease classification, including hypertension [9].

The SVM-PSO hybrid approach combines SVM's strength in handling high-dimensional, non-linear data with PSO's capability in parameter optimization, resulting in better classification accuracy. For instance, [10] applied PSO-SVM to predict hypertension risk in older adults, achieving 93.9% sensitivity, an F1-score of 0.838, and an AUC of 0.871, outperforming conventional SVM.

This study proposes a hybrid classification model combining Support Vector Machine (SVM) and Particle Swarm Optimization (PSO) to enhance diagnostic accuracy in classifying hypertensive patients. By optimizing critical parameters of SVM, the model aims to support faster and more accurate medical decision-making using patient data from RSU at Medika Palopo.

2. Theory

Hypertension is often referred to as the "*Silent Killer*" because people often do not realize they have it until they undergo a health checkup. Accurate and timely classification is essential to prevent severe complications such as stroke and heart failure. Traditional diagnostic methods rely heavily on clinical observations and laboratory testing, which can be time-consuming and resource-intensive. Recent advances in artificial intelligence have enabled the integration of machine learning approaches to improve the precision and efficiency of medical classification systems. This study explores a hybrid classification model that combines SVM and PSO to enhance the predictive performance for hypertension cases. The following section outlines the theoretical foundations that support the methodology applied in this research.

2.1 Medical records

Medical records are documents that contain important information about patients, including their identity, reasons for treatment, time, and methods of care provided during treatment. The main purpose of medical records is to chronologically record the progression of diseases, medical services, and actions taken, so that they can also be used as reference materials for teaching and research in the medical field [11]. The functions of medical records are summarized as ALFRED-AIR, which includes Administration, Legal, Finance, Research, Education, Documentation, Accuracy, Information, and Responsibility [12].

2.2 Digital Image and Image Processing

Hypertension, according to the American Society of Hypertension, is defined as a set of symptoms or syndromes that develop progressively due to other interrelated and complex conditions. Meanwhile, JNC VII states that hypertension occurs when blood pressure exceeds 140/90 mmHg [13]. The mechanism causing hypertension involves the process of forming angiotensin II from angiotensin I through the Angiotensin I Converting Enzyme (ACEI) [14].

2.3 *Artificial Intelligence*

Artificial Intelligence (AI) is a branch of computer science that focuses on developing systems capable of performing tasks that typically require human intelligence, such as decision-making, pattern recognition, and learning from data [15]. AI encompasses various techniques that enable machines to mimic human cognitive functions, including learning, reasoning, and adaptation to the environment [16]

2.4 *Machine Learning*

Machine learning is a mathematical model used to solve specific tasks, particularly in detecting patterns within large data sets. In the field of data science, machine learning is applied to various activities, such as data clustering, forecasting, classification, and so on. Overall, machine learning approaches can be categorized into three types: supervised, unsupervised, and reinforcement [17]. One method in supervised learning is through classification in machine learning. Classification is a type of directed learning that aims to predict target features (which are categorical) for test data based on information obtained from training data [18].

2.5 *Support Vector Machine*

Support Vector Machine (SVM) is an approach used to classify data into several categories. This classification is done by creating a clear dividing line that distinguishes one group from another. Typically, these separating lines are linear. Based on this fundamental principle, SVM is often applied in classification processes, and it falls under supervised learning. Therefore, there are two types of data: training data and testing data. SVM is one of the most widely used techniques in machine learning for both classification and regression purposes [18]. SVM can be used to solve high-dimensional data problems and small training samples. SVM is a method that works based on the principle of Structural Risk Minimization (SRM). SRM is used to maximize the margin and minimize the upper limit of the expected risk from risk [24].

2.6 *Particle Swarm Optimization*

PSO is a computational method used to achieve optimization (obtaining the highest or lowest global value) in certain situations. This process takes place through iteration (repetition) to continuously search for better solutions. The concept of PSO was first introduced by Eberhart and Kennedy in 1995. PSO performs optimization by involving a population consisting of the best candidate solutions and continuously updating these solutions based on specific mathematical formulas [19]. PSO is used to find the optimal SVM parameters (such as C and gamma) to improve the accuracy of diabetes classification. The C parameter balances the trade-off between maximum margin and classification error, while gamma determines how much influence a single data point has on the formation of the decision boundary. PSO operates by forming a population of particles, where PSO works by creating a group of solutions called particles. Each particle moves through the search space to find the best solution based on both individual and collective experience [25].

2.7 *Cross validation*

In the model formation process, random selection performed on the dataset used as training data may not select the most optimal samples for the model. This is because randomly selected samples are not always random, especially when using a small dataset. Cross-validation performs validation by applying test data to a classification model built from random sample selection performed several times [26]. K-fold is one technique of cross-validation. The concept of k-fold cross-validation not only repeats several test data samples, but also divides the dataset into equal parts. The use of k-fold cross-validation can reduce the computation time that may be required due to the iteration process without sacrificing the accuracy of the model estimation [27].

2.8 *Evaluation Metrics*

The Confusion Matrix presents the results obtained during training and testing. The Confusion Matrix also provides an evaluation of classification performance based on objects that are classified correctly or incorrectly. The Confusion Matrix includes actual and predicted information in the classification system. It records prediction results as true positive (TP), true negative (TN), false positive (FP), and false negative (FN), which are then used to compute performance metrics. Equations in the Confusion Matrix model [5], as equation 1. Moreover, Accuracy is the ratio of total correct predictions.:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad [1]$$

Precision measures the proportion of correctly predicted positives, as shown in equation 2.

$$Precision = \frac{TP}{TP+FP} \quad [2]$$

Recall is used to compare the TP ratio to positive events, as shown in equation 3.

$$Recall = \frac{TP}{TP+FN} \quad [3]$$

F1-score is the harmonic mean of precision and recall, as shown in equation 4.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad [4]$$

AUC-ROC is the area under the ROC curve, which describes the model's ability to distinguish between two classes comprehensively by comparing the True Positive Rate (TPR) and False Positive Rate (FPR).

2.9 *Related Research*

Several previous studies have been conducted to examine the application of SVM, PSO, and hybrid approaches in disease classification. These studies serve as important references in designing the classification model proposed in this study, in terms of similarities in methodological approaches, algorithms used, and characteristics of the data analyzed. Several previous studies relevant to the focus of this study are shown in Table 1.

Table 1. Comparison of previous research and recent research

Previous Research	Objective	Methods	Results	Recent Research
Machine Learning in Hypertension Detection: A Study on World Hypertension Day Data [20]	Evaluating the performance of several ML algorithms in hypertension classification	SVM, RF, DT, LR, and XGBoost on original data, oversampling, and undersampling	SVM has high sensitivity (0.790) but low accuracy (0.547) and low AUC (0.619)	This study continues the exploration of SVM with parameter optimization using PSO to improve classification accuracy.
Classification of Hypertension Using the SVM Grid Search Method and SVM Genetic Algorithm (GA) [21]	Evaluating the accuracy comparison of SVM with parameter tuning using Grid Search and Genetic Algorithm	SVM-GS and SVM-GA, RBF and linear kernels	The best accuracy was obtained from SVM-GA (89.71%), but the tuning process was not efficient.	This study proposes the SVM-PSO method for more efficient and accurate parameter optimization in hypertension classification.
Comparison of Support Vector Machine (SVM) and Naïve Bayes Algorithms in Diabetes Disease Classification [22]	Comparing the performance of SVM and Naïve Bayes in diabetes disease classification	SVM (default), evaluated with data split and 10-fold cross-validation	SVM accuracy: 77% (split), decreased to 71% (CV)	This study encourages the exploration of SVM parameter optimization to achieve more stable and accurate performance.

3. Method

This study was conducted from May to August 2025 at Medika Palopo General Hospital. This study aims to evaluate the performance of the SVM algorithm and to develop a hybrid classification model by integrating PSO to improve the predictive performance of SVM.

3.1 Data Collection and Preprocessing

The data used in this study are secondary data obtained from At Medika Palopo General Hospital. The dataset collected consists of 400 samples. Each sample has 12 attributes, namely age, gender, family history of hypertension, systolic blood pressure, diastolic blood pressure, weight, smoking status, eating habits, and stress levels. The preprocessing stage of hypertension data involves processing raw data into a more structured and understandable form before it is used in further analysis. This preprocessing is important for improving data quality and ensuring more accurate model results. The initial stage of data processing in this study consists of cleaning, transforming, and standardizing the data. All data in text format is converted to binary format. Finally, data standardization was performed using the min-max scaling method to equalize the scales between variables, so that no feature would be more dominant in the machine learning process.

3.2 System Design

A Multilayer Perceptron (MLP) architecture was System design is a very important process in software or information system development. This process involves several steps and components that must be considered to ensure that the system built can meet user needs and function properly. Figure 1 illustrates the workflow of the hypertension classification system designed by integrating the PSO dan SVM.

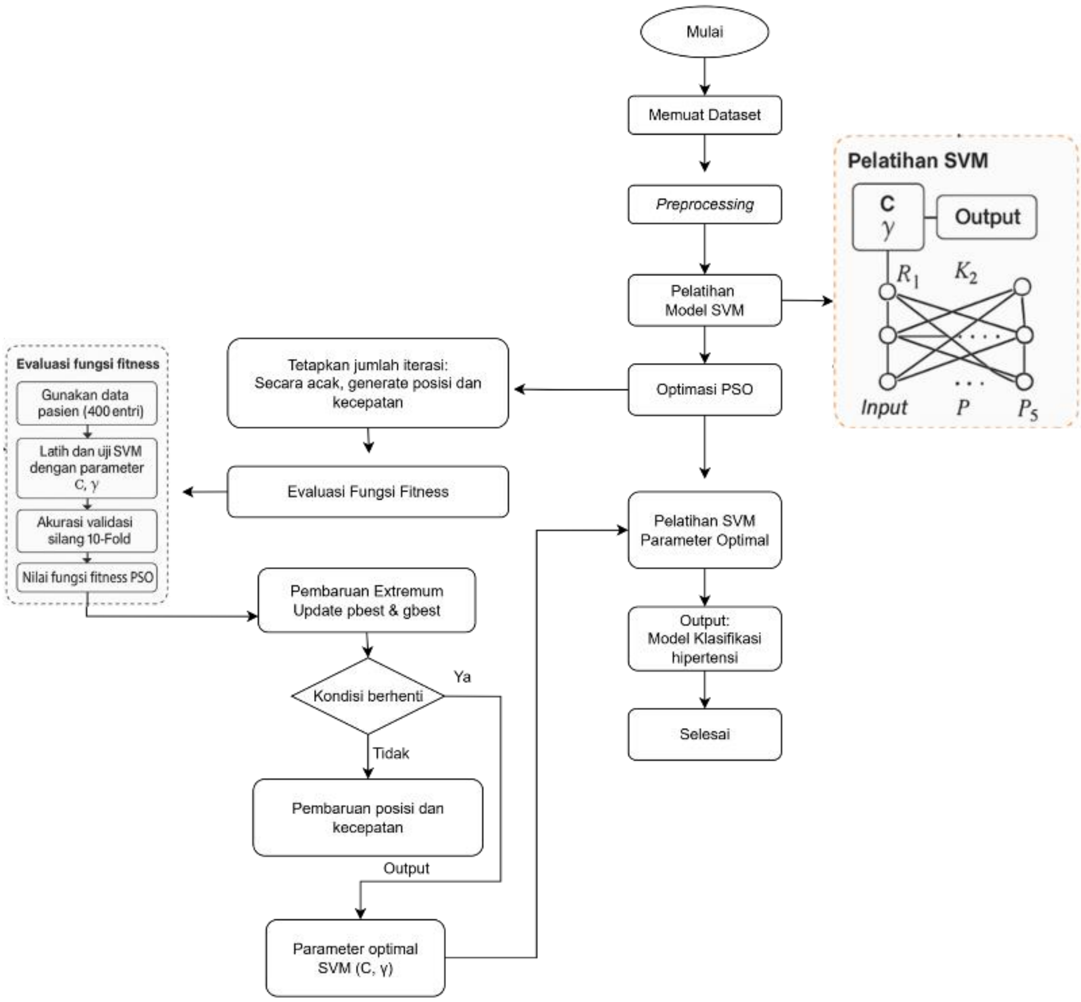

Figure 1. *System Design Flowchart*

The system design begins with the initialization process of the PSO algorithm, where several particles are randomly generated to represent the C and γ parameter values of the SVM. Next, at each iteration, the fitness function is evaluated using 400 patient data entries. The C and γ parameters of each particle are used to train and test the SVM model, and the performance of each parameter combination is measured using the accuracy of 10-fold cross-validation. This accuracy value is then used as the fitness function value that determines the quality of each particle's solution.

If the stopping condition is not met, the process continues by updating the position and velocity of the particles based on the best individual and global experiences. Each particle updates its velocity based on the best experience (pbest) and the best particle in the population (gbest) using the formula [29], as the Equation 5.

$$vi(t+1) = w \cdot v\_i(t) + c1 \cdot r1 \cdot (pbest\_i - x\_i(t)) + c2 \cdot r2 \cdot (gbest - x\_i(t)) \quad (5)$$

Position is updated as equation 6.

$$xi(t+1) = x\_i(t) + v\_i\,(t+1) \quad (6)$$

Notes:

v_i : velocity of particle i
x_i : position (values of parameters C and γ)
w : inertia factor
c1, c2 : learning coefficients
r1, r2 : random numbers between 0 and 1

This process repeats until the maximum iteration is reached or the solution converges. After the optimization process is complete, the optimal SVM parameters (C and γ) that yield the best classification performance are obtained. These parameters are then used for the final SVM training, which subsequently produces a classification model to detect whether a patient falls into the hypertension category or not. This SVM model will be used to automatically and accurately classify new patient data based on patterns learned from training data. The classification function used by SVM is[30], as shown in equation 7.

$$f(x) = sign(\Sigma\ \alpha\_i\ y\_i\ K(x\_i, x) + b) \quad (7)$$

Explanation:

f(x) : decision function to determine the class of new data x
α_i : Lagrange multiplier weights resulting from SVM training
y_i : class labels from training data (e.g., -1 or +1)
K(x_i, x) : kernel function between training data x_i and new data x
b : bias or intercept resulting from SVM training

With a sigmoid kernel, as the equation 8.

$$K(x\_i, x\_j) = tanh(\gamma \cdot x\_i^T\, x\_j + r) \quad (8)$$

Notes:

Tanh : hyperbolic tangent function
γ : kernel parameter that controls the shape of the curve
$x\_i^T x\_j$: dot product between two feature vectors
r : bias constant in the kernel function (specified manually)

3.3 System Testing Techniques

Cross-validation is an important technique in evaluating predictive models to avoid overfitting and improve generalization. In the 10-Fold Cross Validation method, the data is randomly divided into 10 relatively balanced parts. Each portion is sequentially used as test data, while the remaining nine portions are used for model training. This process is repeated 10 times, and the evaluation results from each iteration are averaged to obtain a more stable and reliable measure of model performance. Additionally, a comparison method is used between the conventional SVM, which uses all features without feature selection, and parameters set to default values or manually tuned as a baseline, while the SVM-PSO uses the PSO algorithm to find the optimal values for the C and γ parameters in the SVM, while also performing relevant feature selection. The goal of this approach is to improve the accuracy and efficiency of the model.

3.4 Evaluation

Performance evaluation is carried out using several classification metrics, including accuracy, precision, recall, F1 score, and AUC-ROC.

3.5 Data Analysis Techniques

Data analysis is performed after all stages of training and model testing are complete. The main objective of this stage is to evaluate the effectiveness of the classification approach and interpret the results to support decision-making or further system development.

4. Result and Analysis

The analysis stage evaluates the accuracy and effectiveness of the proposed model in classifying hypertension using the SVM-PSO hybrid method. Several steps were taken to achieve the final model, starting from data collection, preprocessing, implementation, training, optimization, testing, and visualization of results.

4.1 Classification Performance SVM Baseline

At this stage, an initial classification model was built using the SVM algorithm with a sigmoid kernel without parameter optimization. The purpose of this implementation was to form a baseline model, which could then be compared with the optimized model using PSO. The sigmoid kernel was chosen because it is capable of handling nonlinear data with complex characteristics and is relevant for mapping data into higher-dimensional space. The default parameters from the sklearn library are used, including C = 1.0 and gamma = auto. The model is trained using 80% of the data and tested on the remaining 20%.

The analysis evaluates the performance of the baseline Support Vector Machine (SVM) model using a sigmoid kernel for hypertension classification. Table 1 presents key evaluation metrics, including accuracy, precision, recall, F1-score (expressed in percentages), and the area under the curve (AUC). The model achieved an accuracy of 61.25%, precision of 63.22%, recall of 55.60%, F1-score of 59.31%, and an AUC of 0.6400.

Table 2. *Classification Report for SVM*

Methods	Accuracy	Precision	Recall	F1-Score	AUC
SVM (Sigmoid Kernel)	0.6125	0.6322	0.5560	0.5931	0.6400

These results suggest that the model performs moderately well in terms of precision, but its relatively low recall indicates that many hypertensive cases remain undetected. The F1-score, which balances both metrics, confirms the need for improvement in classification consistency. Furthermore, the AUC value of 0.6400 reflects limited discriminatory ability.

In summary, the SVM model without parameter optimization provides only modest classification capability, reinforcing the importance of further enhancement, such as hyperparameter tuning or hybrid approaches to achieve clinically reliable prediction.

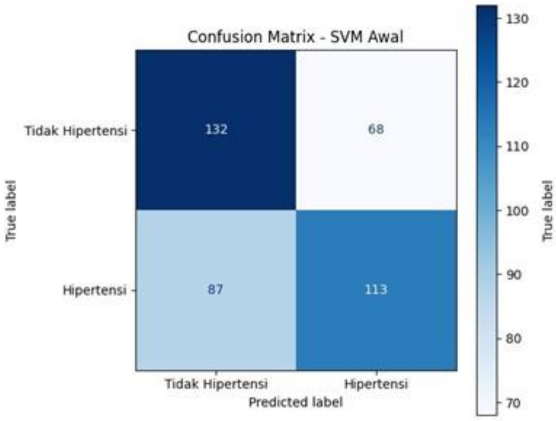

Figure 2. *Confusion Matrix for SVM Baseline*

The confusion matrix in Figure 2 illustrates the classification results of the initial SVM model (without PSO optimization). The model correctly classified 132 non-hypertensive samples, but incorrectly classified 68 samples as hypertensive. For the hypertensive class, 113 samples were correctly classified, while 87 samples were incorrectly classified as non-hypertensive. Based on Figure 2, the accuracy, precision, recall, and F1-score values for the SVM test can be calculated.

Table 3. Confusion Matrix Evaluation Results of SVM

SVM				
TP	**FP**	**TN**	**FN**	**Total Data**
113	68	132	87	400

1. *Calculating Accuracy*

   = (TP + TN) / (TP + TN + FP + FN)

   = (113 + 132) / (113 + 132 + 68 + 87)

   = **0.6125**

2. *Calculating Precision*

   = TP / (FP + TP)

   = 113/(68 + 113)

   = **0.6322**

3. *Calculating Recall*

   = TP / (TP + FN)

   = 113 / (113 + 87)

   = **0.5650**

4. *Calculating F1 Score*

   = 2 * (Precision * Recall) / (Precision + Recall)

   = 2 * (0.6322 * 0.5650) / (0.6322 + 0.5650)

   = **0.5931**

4.2 *Optimization of SVM Parameters Using PSO*

To improve the performance of the SVM classification model, which previously showed moderate accuracy in its initial implementation, a parameter optimization strategy was carried out using the PSO algorithm. PSO was chosen because of its superior ability to explore a wide and complex solution search space, especially in non-convex optimization problems such as SVM kernel parameter tuning. In the context of the SVM algorithm with a sigmoid kernel, two main parameters significantly influence classification performance.

1) *C (Regularization Parameter),* which is a parameter that regulates the balance between maximizing the separating margin and minimizing classification errors. C value that is too large can cause overfitting because the model will try hard to classify all data correctly. Conversely, a value of C that is too small will result in a wide margin but carries a high risk of underfitting.
2) *Gamma (γ)* determines the influence of a single data point on the formation of the hyperplane. A high gamma value can cause the model to focus too much on specific data (overfitting), while a gamma value that is too small makes the model less sensitive to patterns in the data generated.

The optimization process employed k-fold cross-validation to evaluate each parameter combination generated by PSO, ensuring more accurate and reliable performance estimates by alternately using all data as training and testing sets. This method reduces evaluation bias and prevents overfitting, ensuring the selected parameters yield a model that generalizes well to new data. Table 4 presents the PSO-based parameter optimization process for SVM, involving the determination of population size, maximum iterations, initialization ranges for C and gamma, and computation time.

Table 4. Particle Swarm Optimization Process

Parameters Optimization				
Population Size	**Maximum Iterations**	**Initialization Range**		**Computation time**
		C	***Gamma***	
20	10	0,1-100	0,0001-10	± 30 seconds

Table 5. Particle Swarm Optimization Results

Optimization	C	Gamma
PSO	27.3927	0.0320

Each particle represents a pair of values (C, gamma) and moves within the solution space by considering both personal and global best positions. The process iterates until the maximum iteration is reached or no significant improvement in the fitness function is observed. The fitness function, defined as classification accuracy on the training data, reflects the model's ability to correctly classify instances. After 10 iterations, the optimization identified the best parameter combination, as shown in Table 5.

4.3 Classification Performance of SVM-PSO

After the parameter optimization process was carried out using the PSO algorithm, the optimal values for the C and gamma parameters were C = 27.3927and gamma =0.0320. These values were then used to retrain the SVM model with a sigmoid kernel. The purpose of this implementation was to evaluate the extent of the improvement in classification performance achieved by the hybrid SVM-PSO approach compared to the previous baseline model.

The analysis evaluates the performance of the SVM model after parameter optimization using PSO for hypertension classification. Table 2 presents the classification metrics expressed in percentages, including accuracy, precision, recall, F1-score, and the area under the curve (AUC).

The optimized model achieved an accuracy of 96.75%, precision of 97.14%, recall of 96.31%, F1-score of 96.73%, and an AUC of 0.9725. These results indicate a substantial improvement over the baseline SVM model, with consistently high values across all evaluation metrics. The near-perfect precision and recall demonstrate the model's ability to correctly identify both hypertensive and non-hypertensive cases with minimal error.

These results suggest that the hybrid SVM-PSO model performs exceptionally well across all metrics. The high precision of 97.14% indicates the model's strong ability to correctly identify positive hypertension cases with minimal false positives. Similarly, the recall of 96.31% reflects a low rate of false negatives, meaning most hypertensive cases were successfully detected. The F1-score of 96.73% confirms a well-balanced performance between precision and recall. Furthermore, the AUC value of 0.9725 demonstrates excellent discriminatory ability in differentiating between hypertensive and non-hypertensive patients. In summary, the SVM model optimized with PSO provides highly accurate and reliable classification, highlighting the effectiveness of the hybrid approach in clinical prediction tasks.

Table 3. *Classification Report Of SVM-PSO*

Methods	Accuracy	Precision	Recall	F1-Score	AUC
SVM + PSO	0.9675	0.9714	0.9631	0.9673	0.9725

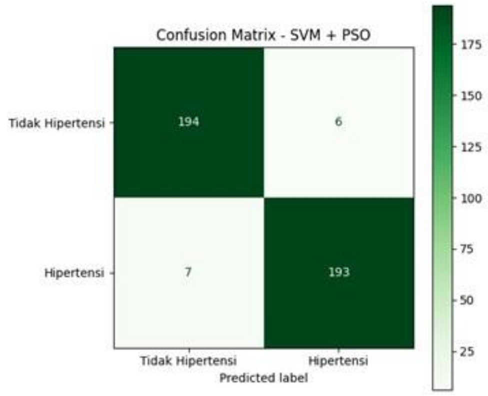

Figure 3. *Confusion Matrix for SVM-PSO*

The confusion matrix in Figure 3 shows an improvement in model performance after parameter optimization using the PSO algorithm. The model successfully classified 194 non-hypertensive samples and 193 hypertensive samples correctly. There were only 6 non-hypertensive samples and 7 hypertensive samples that were misclassified. Based on Figure 2, the accuracy, precision, recall, and F1-Score values for the SVM-PSO test can be calculated.

Table 6. Confusion Matrix Evaluation Results of SVM-PSO

SVM				
TP	**FP**	**TN**	**FN**	**Total Data**
113	68	132	87	400

1. *Calculation Accuracy*

   = (TP + TN) / (TP + TN + FP + FN)

   = (193 + 194) / (193 + 193 + 6 + 7)

   = **0.9675**

2. *Calculation Precision*

   = TP / (FP + TP)

   = 193 / (6 + 193)

   = **0.9714**

3. *Calculation Recall*

   = TP / (TP + FN)

   = 193 / (193 + 7)

   = **0.9650**

4. *Calculation F1 Score*

   = 2 * (Precision * Recall) / (Precision + Recall)

   = 2 * (0.9714 * 0.9650) / (0.9714 + 0.9650)

   = **0.9673**

4.4 Comparison of Test Results

4.4.1 Comparison of the Impact of SVM Parameters on Performance

Parameter optimization is a crucial component in improving the performance of machine learning algorithms. Table 7 shows the differences in the main parameter values in SVM before and after the PSO optimization.

Before optimization, the values of C and gamma are set by default by the Scikit-learn library, namely C = 1.0 and gamma = 'scale', which are not always suitable for the characteristics of the dataset. After optimization, the C value increases to 27.3927, indicating that the model is given more flexibility to minimize classification errors. Meanwhile, the gamma value changes to 0.0320, providing a broader range of influence from each data point on the model's decision.

Table 7. Comparison of Parameter Impacts on Performance

Parameters	**Before Optimization**	**After Optimization**
Kernel	Sigmoid	Sigmoid
C	1.0	27.3927
Gamma	Scale (automatic)	0.0320

4.4.2 Comparison of SVM and SVM-PSO Evaluation Metrics Against Test Results

Table 8 presents a comparison of evaluation metrics derived from the confusion matrix for both models, SVM and PSO optimized SVM (SVM+PSO). The comparison highlights performance improvements after optimization by examining True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN).

Table 8. *Comparison of Evaluation Metrics*

SVM				**SVM+PSO**			
TP	FP	TN	FN	TP	FP	TN	FN
113	68	132	87	193	6	194	7

Table 8 shows that SVM-PSO reduced false positives from 68 to 6 and false negatives from 87 to 7, indicating significantly higher accuracy in correctly identifying both hypertensive and non-hypertensive patients, with fewer classification errors.

4.4.3 Comparison of SVM and SVM-PSO Test Results

Furthermore, Table 9 compares the key evaluation metrics—accuracy, precision, recall, F1-score, and AUC—between the SVM model and the PSO-optimized SVM (SVM+PSO). These metrics collectively measure the classification performance of each model. The comparison illustrates the performance improvements achieved after applying PSO optimization, as detailed in Table 9.

Table 9. *Comparison of Results*

No	Methods	Accuracy	Precision	Recall	F1-score	AUC
1	SVM	61,25%	63,22%	56,50%	59,31%	0,6400
2	SVM+PSO	96,75%	97,14%	96,50%	96,73%	0,9732

These differences indicate that applying parameter optimization using PSO leads to a substantial improvement in model performance across all metrics:

a) Accuracy increased to 94.99%, indicating that nearly all test data, both hypertensive and non-hypertensive patients, were correctly classified.
b) A precision of 94.84% shows that most predicted hypertensive cases were indeed correct, resulting in a very low false positive rate.
c) Recall of 95.45% reflects the model's ability to detect almost all hypertensive patients, minimizing false negatives.
d) A high F1-score (95%) demonstrates an excellent balance between precision and recall, indicating no bias toward either class.
e) An AUC of 0.9574 indicates a strong discriminative ability to distinguish between hypertensive and non-hypertensive patients, as shown by the ROC curve approaching the top-left corner.

These findings validate that the hybrid SVM-PSO approach significantly outperforms the conventional SVM model without optimization.

4.4.4 Accuracy Comparison of SVM-PSO Across Iterations

To evaluate the effectiveness of the PSO optimization process, model accuracy was observed over varying iteration counts. This analysis aimed to assess the stability of performance improvements and identify the optimal point at which accuracy converges. The detailed accuracy comparison for each iteration is presented in Table 10.

Table 10. Accuracy Comparison Across Iterations

Iterations	Accuracy
1	89 %
2	90,20 %
4	90,25 %
6	96,00%
8	96,50 %
10	96,75%

From Table 10, it can be seen that the model's accuracy consistently improved from the initial iterations, starting at 89.00% in the first iteration, increasing to 90.25% by the fourth, and reaching its optimal value of 96.75% at the tenth iteration. These results indicate that the PSO algorithm efficiently explores the solution space, achieving optimal performance without requiring excessive iterations. The stable accuracy at iteration 10 also suggests that the optimization process did not lead to overfitting but effectively refined the model parameters.

4.5 AUC-ROC Curve Analysis Between SVM and SVM-PSO

The AUC-ROC curve in Figure 4 shows a significant improvement in the model's discriminative ability after PSO optimization. The pre-optimization SVM achieved an AUC of 0.64, indicating limited classification capability, whereas the PSO-optimized model reached an AUC of 0.97, categorized as excellent. This demonstrates that the optimized model can accurately distinguish between hypertensive and non-hypertensive patients. The performance gain confirms that PSO successfully identified the optimal parameters for the SVM sigmoid kernel, resulting in more accurate classification. Similar findings have been reported in previous studies, where PSO-SVM achieved high AUC values after preprocessing and data balancing in stroke classification tasks.

Figure 4 illustrates the AUC-ROC curves of the SVM model before and after PSO optimization. The blue line represents the pre-optimization SVM, with an AUC of 0.64, which is close to the diagonal reference line (Random Classifier), indicating limited discriminative ability. In contrast, the orange curve (SVM+PSO) shows a sharp increase in the True Positive Rate (TPR) with a low False Positive Rate (FPR). The curve's proximity to the top-left corner reflects excellent classification performance, with an AUC of 0.97. An AUC close to 1 indicates a very low error rate, confirming that PSO parameter optimization significantly enhanced the model's sensitivity and specificity, making it more reliable for predicting hypertension in the given dataset.

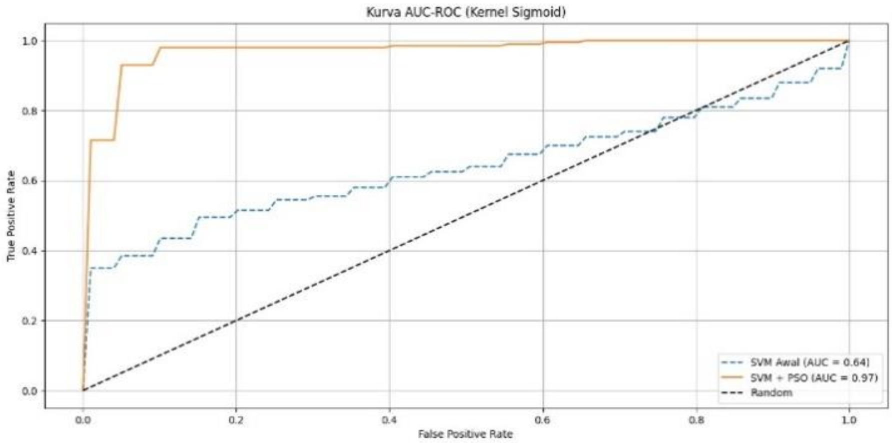

Figure 4. Curve AUC-ROC of SVM and SVM-PSO

4.6 Discussion and Comparison

The proposed hybrid SVM-PSO method demonstrates competitive performance compared to [23], who applied SVM-PSO with SMOTE for stroke classification. Despite not using SMOTE, the model in this study achieved comparable results, highlighting the effectiveness of proper feature selection, preprocessing, and parameter tuning. Furthermore, this study presents a wider and explicitly defined search range for PSO optimization, adding novelty in the optimization approach and supporting efficient classification with reduced algorithmic complexity.

4.7 Summary

In summary, the application of the hybrid SVM-PSO approach has been proven to significantly improve classification performance. After optimizing the sigmoid kernel parameters using the PSO algorithm, the accuracy value increased to 94.75%, with an AUC of 0.9555. In addition, precision and recall also increased to nearly 95%. This proves that hyperparameter optimization using PSO is effective in improving the generalization ability of the SVM model.

4.8 Recommendations

Future research could try applying this to a wider and more diverse dataset, exploring other optimization methods, using more comprehensive additional features, and classifying based on the severity of hypertension.

5. Conclusions

This study demonstrates that the hybrid SVM-PSO approach significantly enhances the classification performance of hypertension diagnosis compared to conventional SVM. Using 400 patient records with 12 clinical features from At Medika Palopo General Hospital, the hybrid model optimized key parameters (C and gamma) via PSO, resulting in a substantial performance improvement—from 61.25% to 96.75% in accuracy and from 0.6400 to 0.9725 in AUC. These results highlight the potential of SVM-PSO as an effective AI-based medical decision support tool for early detection of hypertension, which is crucial to prevent severe complications such as stroke and heart failure. Future work may explore its application to classify hypertension severity levels and validate its generalizability across larger and more diverse clinical datasets.

Acknowledgments: The author would like to thank Muhammadiyah University Makassar, particularly the Department of Informatics, for its support and facilities provided during the research. Thanks are also extended to Medika Palopo General Hospital and all parties who contributed to data collection and system testing.

Author contributions: The authors are responsible for building Conceptualization, Methodology, analysis, and investigation: **Utami, R., Anggreani, D., & Darniati.** Data curation, writing—original draft preparation, writing—review and editing, visualization: **Utami, R., Anggreani, D., & Darniati.** Supervision of project administration, funding acquisition: **Utami, R., Anggreani, D., & Darniati.** and have read and agreed to the published version of the manuscript.

Funding: The study was conducted without any financial support from external sources.
Availability of data and Materials: All data are available from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Additional Information: No Additional Information from the authors.

References

[1] R. Sarfika and S. I Made Moh. Yanuar, Perawatan Diri Penderita Hipertensi Usia Dewasa: Berbasis Teori dan Riset. Jakarta: Deepublish, 2024. [Online]. Available: https://books.google.co.id/books?id=AhwZEQAAQBAJ
[2] H. Akbar, T. M. Rafsanjani, A. H. Sinaga, W. R. Hidayani, Y. Panma, and S. R. Bela, Teori Epidemiologi Penyakit Tidak Menular. Aceh: Yayasan Penerbit Muhammad Zaini, 2021. [Online]. Available: https://books.google.co.id/books?id=FmBQEAAAQBAJ
[3] N. Novianti, S. P. A. Alkadri, and I. Fakhruzi, "Klasifikasi Penyakit Hipertensi Menggunakan Metode Random Forest," Progresif J. Ilm. Komput., vol. 20, no. 1, p. 380, 2024, doi: 10.35889/progresif.v20i1.1663.
[4] M. D. F. Tino, Herliyani Hasanah, and Tri Djoko Santosa, "Perbandingan Algoritma Support Vector Machines (SVM) Dan Neural Network Untuk Klasifikasi Penyakit Jantung," INFOTECH J., vol. 9, no. 1, pp. 232–235, 2023, doi: 10.31949/infotech.v9i1.5432.

[5] W. S. Dharmawan, "Komparasi Algoritma Klasifikasi Svm-Pso Dan C4.5-Pso Dalam Prediksi Penyakit Jantung," I N F O R M a T I K a, vol. 13, no. 2, p. 31, 2022, doi: 10.36723/juri.v13i2.301.
[6] N. Cao and W. Wang, "Research on SVM Parameter Optimization Mechanism Based on Particle Swarm Optimization," 2021, pp. 328–337. doi: 10.1007/978-3-030-78615-1_29.
[7] S. M. Almufti, A. Yahya Zebari, and H. Khalid Omer, "A comparative study of particle swarm optimization and genetic algorithm," J. Adv. Comput. Sci. Technol., vol. 8, no. 2, pp. 40–45, 2021, doi: 10.14419/jacst.v8i2.29401.
[8] E. Purwaningsih and E. Nurelasari, "Peningkatan Akurasi Metode Support Vector Machine melalui Particle Swarm Optimization pada Penyakit Ginjal Kronis," Inf. Manag. Educ. Prof., vol. 9, no. 1, pp. 61–70, 2024.
[9] R. You et al., "Development and Validation of a Hypertension Risk Prediction Model Based on Particle Swarm Optimization–Support Vector Machine," Bioengineering, vol. 12, no. 3, pp. 1–19, 2025, doi: 10.3390/bioengineering12030238.
[10] S. K. M. M. P. H. Feby Erawantini, M. K. dr. Arinda Lironika Suryana, and S. K. M. K. Khoirunnisa' Afandi, Rekam Kesehatan Elektronik Dengan Clinical Decision Support System (CDSS). Erawantini: UPT Penerbitan & Percetakan Universitas Jember, 2021. [Online]. Available: https://books.google.co.id/books?id=b9xKEAAAQBAJ
[11] S. Anjani and M. T. Abiyasa, Disrupsi Digital dan Masa Depan Rekam Medis (Kajian Peraturan Menteri Kesehatan Nomor 24 Tahun 2022 Tentang Rekam Medis Elektronik). Yogyakarta: Selat Media, 2023. [Online]. Available: https://books.google.co.id/books?id=aDvJEAAAQBAJ
[12] D. Dwisari, Buku Ajar Patofisiologi edisi 2. Indramayu: Penerbit Adab, 2024. [Online]. Available: https://books.google.co.id/books?id=guQKEQAAQBAJ
[13] Y. Ardilla et al., DATA MINING DAN APLIKASINYA. Bandung: Penerbit Widina, 2021. [Online]. Available: https://books.google.co.id/books?id=53FXEAAAQBAJ
[14] N. M. Arhami, Muhammad, Data Mining - Algoritma dan Implementasi. Lhoksumawe: Penerbit Andi, 2022.
[15] C. C. Aggarwal, Artificial Intelligence: A Textbook. Switzerland: Springer US, 2021.
[16] V. Y. P. Ardhana et al., Konsep Dasar Teknologi Informasi. Sumedang: MEGA PRESS NUSANTARA, 2024. [Online]. Available: https://books.google.co.id/books?id=siceEQAAQBAJ
[17] B. A. S. Fairuz, Panduan Praktis Machine Learning Klasifikasi Menggunakan Python: Diandra Kreatif. Yogyakarta: Diandra Kreatif, 2024. [Online]. Available: https://books.google.co.id/books?id=Wr5AEQAAQBAJ
[18] D. Wulandari, S. Aziz, S. Adrianto, F. Pratiwi, and R. M. Sari, Teori Dan Implementasi Machine Learning Menggunakan Python. Payakumbuh: Serasi Media Teknologi, 2025. [Online]. Available: https://books.google.co.id/books?id=GNJNEQAAQBAJ
[19] Nuryani and F. Alhafid, Implementasi Swarm Support Vector Mechine untuk Deteksi Fibrilasi Atrium Menggunakan Elektrokardiogram. Jakad Media Publishing, 2024. [Online]. Available: https://books.google.co.id/books?id=c2cQEQAAQBAJ
[20] S. Montagna et al., "Machine Learning in Hypertension Detection: A Study on World Hypertension Day Data," J. Med. Syst., vol. 47, no. 1, pp. 1–10, 2023, doi: 10.1007/s10916-022-01900-5.
[21] F. O. Awalullaili, D. Ispriyanti, and T. Widiharih, "Klasifikasi Penyakit Hipertensi Menggunakan Metode Svm Grid Search Dan Svm Genetic Algorithm (Ga)," J. Gaussian, vol. 11, no. 4, pp. 488–498, 2023, doi: 10.14710/j.gauss.11.4.488-498.
[22] A. Desiani, N. R. Dewi, M. Arhami, D. S. Sitorus, and S. Rahmadita, "Comparison of Support Vector Machine (SVM) and Naïve Bayes Algorithms in Diabetes Disease Classification," POSITIF J. Sist. dan Teknol. Inf., vol. 10, no. 1, pp. 65–74, 2024.
[23] Y. Ayuningtyas and I. Made Suartana, "Klasifikasi Penyakit Stroke Menggunakan Support Vector Machine (SVM) dan Particle Swarm Optimization (PSO)," JINACS (Journal Informatics Comput. Sci., vol. 04, no. 04, pp. 452–457, 2023.
[24] S. Kenny Riva, S. Casi, and S. Randy Erfa, "Analisis Algoritma Support Vector Machine Dalam Klasifikasi Penyakit Stroke," e-Proceeding Eng., vol. 9, no. 3, p. 922, 2022.
[25] A. Agung, G. Agung, and I. M. Widiartha, "Optimasi Metode Support Vector Machine (SVM) Mengunakan Particle Swarm Optimization pada Permasalahan Klasifikasi Diabetes," vol. 3, pp. 879–888, 2025.
[26] A. Maulana, A. Nugroho, and I. Romli, "Optimalisasi Support Vector Machine Menggunakan Particle Swarm Optimization Untuk Mendiagnosa Penyakit Kanker Payudara," J. Pract. Comput. Sci., vol. 1, no. 2, pp. 1–11, 2022, doi: 10.37366/jpcs.v1i2.940.
[27] M. R. Maulana, A. Sucipto, and H. M. Mulyo, "Optimisasi parameter Support Vector Machine dengan Particle Swarm Optimization untuk peningkatan klasifikasi diabetes," J.Inform. Teknol. dan Sains (JINTEKS), vol. 6, no. 4, pp. 802–812, 2024.
[28] D. K. Choubey, S. Tripathi, P. Kumar, V. Shukla, and V. K. Dhandhania, "Classification of Diabetes by Kernel Based SVM with PSO," Recent Adv. Comput. Sci. Commun., vol. 14, no. 4, pp. 1242–1255, 2019, doi: 10.2174/2213275912666190716094836.
[29] W. Huang et al., "Railway dangerous goods transportation system risk identification: Comparisons among SVM, PSO-SVM, GA-SVM, and GS-SVM," Appl. Soft Comput., vol. 109, p. 107541, 2021, doi: 10.1016/j.asoc.2021.107541.
[30] Danis Rifa Nurqotimah, A. Naseh Khudori, and R. Siwi Pradini, "Implementasi Algoritma Support Vector Machine (SVM) Untuk Klasifikasi Penyakit Stroke," J. Appl. Comput. Sci. Technol., vol. 5, no. 2, p. press, 2024, doi: 10.52158/jacost.v5i2.817.